
UPACK Program Package for Crystal Structure Prediction: Force Fields and Crystal Structure Generation for Small Carbohydrate Molecules

BOUKE P. VAN EIJCK, JAN KROON

Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

Received 23 July 1998; accepted 15 January 1999

ABSTRACT: The UPACK program package for crystal structure generation was used to build hypothetical crystal structures for 32 pyranoses and 24 polyalcohols. A subset of these was used to compare six force fields in their ability to reproduce the experimentally observed structures, preferably with low energies and high rankings with respect to the structure with lowest energy. One of these force fields, UNITAT, was derived from GROMOS87 by adjusting some parameters to obtain a better geometric description of these crystal structures. This united-atom force field and the all-atom OPLS force field performed best, and were used in the crystal structure generation for the full set of compounds. For these carbohydrates, hundreds of hypothetical polymorphic structures are generated in an energy window of about 25 kJ/mol. On average, the experimental structures had reasonable energies and rankings. These lists are good starting points for more sophisticated calculations, and can be used for structure determination by powder diffraction methods. In a few cases, serious doubt was raised concerning the hydrogen-bond network reported in the literature, for which more plausible alternatives appear to exist. © 1999 John Wiley & Sons, Inc. J Comput Chem 20: 799–812, 1999

Keywords: crystal structure prediction; force fields; carbohydrates; polymorphism

Introduction

The continuous increase in easily available computer power has allowed the development of several methods to generate crystal structures without use of the experimental data. Reviews on this subject have been done by Gdanitz¹ and Verwer and Leusen.² Most methods rely on energy calculations with empirical force fields. Thus, kinetic and thermal effects are not considered, and no phase transitions between polymorphs can be predicted—in fact, one may doubt whether crystal structures are predictable at all.³ In our opinion, we have not yet reached the stage where the calculated energies are accurate enough to warrant a study of more subtle effects. Possible exceptions occur for very simple molecules like benzene⁴ or sophisticated force fields such as those used by Price.^{5,6}

All experience gained so far indicates that the number of hypothetical structures can be quite large. The energy ordering can be affected by the choice of the force field.⁷ In fact, this provides us with a new and very severe test for a force field: the observed crystal structure should be present on the list of hypothetical crystal structures, not only with a good geometry but also with a reasonably low energy.⁸

Most work done so far concerns structures that are constructed from rigid molecules. Our study on the prediction of hexapyranose structures⁹ is one of the first in which essentially flexible molecules are considered. Here, the torsional angles of the hydroxyl hydrogen atoms about their C—O bonds are completely unpredictable *a priori*, because they are determined mainly by the geometry requirements of the hydrogen-bond network that holds the carbohydrate structure together. In addition, there may occur freedom of internal rotation for heavier groups, which gives rise to a number of distinct conformations whose geometries are fairly easy to model for a free molecule. Each form has to be considered as a possible building block in the crystal structure prediction. We shall reserve the name “conformation” for a structure defined by this type of internal rotation, excluding the different rotamers for the hydroxyl groups.

Thus, it is a difficult task to generate possible crystal structures for carbohydrates with a reasonable confidence that the experimentally observable structure(s) will be among the hypothetical ones.

Presently we have chosen a brute-force grid search as the most reliable method in this respect. For simpler molecules algorithms based on some kind of random search are more efficient, and we are planning to investigate the application of such methods to carbohydrates in the near future.

In this work we first give a short description of the program package UPACK, which performs the generation of hypothetical crystal structures. We then compare the performance of several force fields for carbohydrates, including an improved united-atom force field, and examine the crystal structure generation for 32 pyranoses and 24 polyalcohols. The goal is to reduce the number of possible structures to, say, less than 50, which can be used as starting point for further study with a more sophisticated force field.

UPACK Program Package

The program suite UPACK (Utrecht Crystal Pack-er) is designed to construct crystal structures of low potential energy, using a molecular force field. It is, at present, restricted to crystal structures with one molecule in the asymmetric unit, and to space groups with triclinic, monoclinic, or orthorhombic symmetry. The programs have been designed with simple carbohydrates as target molecules, but should be usable for other chemical classes if one is prepared to extend some default parameter files.

The basic form of the force field is:

$$\begin{aligned}
 U = & \sum_{\text{bonds}} \frac{1}{2} k_b (b - b_0)^2 + \sum_{\text{angles}} \frac{1}{2} k_\theta (\theta - \theta_0)^2 \\
 & + \sum_{\text{impropers}} \frac{1}{2} k_\xi (\xi - \xi_0)^2 \\
 & + \sum_{\text{dihedrals}} \sum_n [|V_n| + V_n \cos(n\omega)] \\
 & + \sum_i \sum_{j>i}' \left[\frac{C_{12}(ij)}{r_{ij}^{12}} - \frac{C_6(ij)}{r_{ij}^6} + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right] \quad (1)
 \end{aligned}$$

Improper dihedral angles are needed to maintain chirality around united atoms, and may be helpful to stabilize pyranose ring conformations. The form of the dihedral potential was chosen to be consistent with other force fields, allowing a negative V_n without using a phase factor of 0 or π in the cosine. The prime in the summation over the non-bonded atoms i and j indicates the omission of interactions between atoms that are connected through one or two bonds.

A rigid molecule is taken as the asymmetric unit in a given space group, and placed with varying positions and orientations into a unit cell of varying dimensions. This gives, at most, 12 ($P\bar{1}$) and, at least, 8 (e.g., $Pna2_1$) parameters to be varied, which can be done either in a systematic grid search or by assigning random values. To avoid wasting time on unreasonable structures, one of the cell parameters can be eliminated by using an estimated density. Each energy calculation is discontinued immediately if any repulsion term exceeds a given value. This will occur in the majority of cases. The remaining ones are subjected to energy minimization, and are only kept if the decreasing energy reaches certain thresholds, which can be set automatically by the program. A separate calculation has to be done for each conformation. Details of the grid search and the coordinate system have been published earlier.^{4,7}

In this first stage, it is necessary to use united-atom OH groups, because nothing is known about the hydroxyl dihedral angles. Here, a very simple force field is needed: speed is essential because millions of trial structures have to be investigated. As in our previous study,⁹ all charges were omitted and the GROMOS87¹⁰ Lennard-Jones potentials were used; the hydroxyl united OH "atoms" have a special interaction potential to mimic hydrogen bonding. The cutoff distance was increased to 8 Å for greater stability. This is obviously an oversimplified force field, but, as long as adequate starting structures for the following stages are obtained, its precise properties are unimportant.

A few thousand low-energy structures are taken over to the second stage, where the hydroxyl torsion angles are examined. Three starting dihedral angles are taken about each C—O bond, and in the energy minimizations only these angles and the crystallographic parameters are optimized. In the end, a full structure relaxation may be performed. The different conformations are studied consecutively, and the results are merged. In this stage the hydroxyl hydrogen atoms are modeled by charges only. Here, a cutoff radius of 7 Å, based on neutral charge groups rather than on individual atoms, was found to be sufficient to obtain stable energy minimizations. At the beginning of this investigation we used the standard GROMOS87 force field. As will be discussed in what follows, this was later replaced by a slightly modified version.

Again, a few thousand low-energy structures are taken over to the third and final stage where the energy minimizations can be continued with a better and more elaborate force field. United CH,

CH₂, and CH₃ "atoms" are usually replaced by explicit carbon and hydrogen atoms. Ewald summation may be used for the Coulomb and attractive r^{-6} dispersion terms, and a Buckingham-type exponential repulsion may be used instead of the r^{-12} term. Atomic polarizabilities may also be included. It is here that the performance of different force fields can be compared. UPACK also provides the possibility of relaxing the space-group symmetry by expanding to $P1$, and to perform a primitive molecular dynamics "shake-up." In this way, the stability of the obtained crystal structures can be investigated.⁷

All energy minimizations are performed by a number of steepest descent steps, followed by a more or less exhaustive variable metric procedure based on the BFGS algorithm [11, p. 308]. A considerable speed-up is obtained by using lists containing information on nonbonded atom pairs and vectors in reciprocal space. These lists are calculated in the beginning, and are only updated upon a significant change of any geometric parameter. If Ewald summation is turned off, these lists are even essential to obtain stable convergence: when a charge group passes the cutoff surface, the energy jumps noticeably, and its derivatives become unreliable. A further acceleration is possible by omitting all terms with a small energy contribution from the lists. For a threshold of 0.02 kJ/mol, a fourfold decrease in computer time was obtained. Several tests indicated that an error of about 1 kJ/mol in the total energy may result, which is less than the accuracy of the force fields used.

Unavoidably, the same structures will be encountered repeatedly, probably with different choices for the unit-cell parameters. In every stage of the calculation it is essential to eliminate such equivalent structures by a clustering procedure¹² before proceeding to the next stage. The algorithm fails to recognize equivalent hydroxyl torsional angles in flexible molecules with a symmetry element, and such structures were eliminated by inspection.

Force Field Comparison

TEST PROCEDURE

Quite a few force fields for carbohydrates have been published. A comparison, based on energy minimization of selected molecular models, has been organized by Perez.¹³ The problem with such

calculations is that hydroxyl group orientations in free molecules are generally artificial, as they are directed by intramolecular hydrogen bonds rather than by intermolecular ones, which are dominant in more realistic environments.

For reliable crystal structure predictions a very sophisticated force field is needed, with at least a good description of electrostatic interactions.⁸ Many derivatives of AMBER^{14–17} require the determination of charges by *ab initio* calculations. Even better is the use of distributed multipole moments.¹⁸ However, it is by no means clear how such calculations should be performed for flexible molecules where the charges depend on the conformation and on the orientations of the hydroxyl groups.

Moreover, for flexible molecules the intramolecular energy must also be well reproduced to find the best balance between conformational energy and intermolecular interactions. Here force fields like MM3¹⁹ or CLASS II²⁰ are required. They have such complex forms that it is not easy to incorporate them into the UPACK energy routine. In this work, we limited ourselves to force fields that have a simple functional form as well as a straightforward recipe for determining partial atomic charges. As stated earlier, the goal is to obtain a set of structures that is fairly certain to contain the experimentally observable structure(s) and yet small enough to be manageable for more sophisticated calculations.

For united-atom force fields we chose the two versions of GROMOS.^{10,21} We did not investigate the GROMOS reparameterization by Ott and Meyer,²² because it contains a torsion potential that is not invariant for coordinate inversion. In the GROMOS96 force field, the old energy expressions for bonds and angles were retained, and only the revised nonbonded potentials were adopted. For all-atom force fields we chose the Ha et al.²³ version of CHARMM, the adaptation by Kouwijzer et al.,²⁴ and the recent carbohydrate OPLS force field.²⁵ We would have liked to test the consistent force field,^{26,27} but we could not reconstruct the charge allocation algorithm.

To judge these force fields by their capability to predict crystal structures, a limited test set of pyranoses and polyalcohols was selected. The structures are generally indicated only by their refcodes in the Cambridge Structural Database.²⁸ In our previous work,⁹ six hexapyranoses were studied, each in three possible conformations for the exocyclic CH₂OH group. In four cases, the experimental structure had a satisfactorily low energy, but in

the two other ones the energy was some 20 kJ/mol higher than for the most favorable hypothetical structure. This set was now extended with four 1-O-methylhexapyranoses, which have two exocyclic rotors and, therefore, nine possible conformations. The ring configuration was always taken to be ⁴C₁. We further studied the three chiral hexitols, mannitol (for which two polymorphs are well characterized), iditol, and glucitol. These substances allow rotation about five single bonds, so there are, in principle, 243 distinct conformations of which we selected the most plausible ones following Jeffrey and Kim.²⁹ In the notation of these investigators the dihedral angles about the three central C—C bonds should be AAA for mannitol, AMA or PAP for iditol, and MAA or APP for glucitol. To the latter we added AAA, which was found to occur in a recently analyzed polymorph³⁰ where it cocrystallizes with a molecule in the MAA form. Together with the conformational freedom about the terminal C—C bonds this gave 9, 18, and 27 conformations for the three compounds, respectively.

An important requirement for a force field is that it should reproduce the observed crystal geometries. This was verified by comparing a set of various geometrical parameters (Table I) before and after energy minimization of the experimental structures. Here, we do not include the hydrogen atoms, which are unreliable in X-ray determinations, except for the C—C—O—H torsional angles, which are too important to neglect. Ideally, each parameter shift (Δ) should be zero; we compare the actually observed shifts with selected acceptable values, σ , and define an average geometrical disagreement index:

$$D = \sqrt{\langle (\Delta/\sigma)^2 \rangle} \quad (2)$$

TABLE I.
Acceptable Geometry Shifts, σ , upon Energy Minimization of Experimental Structures.

<i>B</i>	Bond lengths without H	0.01 Å
<i>A</i>	Angles without H	1.0°
<i>I</i>	Improper dihedral angles	1.0°
<i>T</i>	Torsions without H	1.0°
<i>TH</i>	Torsions with H	10°
<i>HB</i>	O···O distances < 3.5 Å	0.02 Å
<i>C</i>	Cell axes	0.05 Å
<i>G</i>	Molecular centers of gravity	0.02 Å
<i>E</i>	Molecular Euler angles	1.0°

For various parameters the corresponding values of σ were estimated by comparing the results of different determinations of one and the same crystal structure. The experimental parameters vary enormously in quality, and we took conservatively large values for σ (Table I), because our data set included fairly old information.

We tried to improve the GROMOS87 force field in such a way that the average D for all structures was lowered. In particular, we noted that hydrogen bonds were generally too short and so the hydroxyl charges and the oxygen C_6 parameters were decreased. Moreover, parameters for bond lengths and bond angles were changed and the angle force constants were increased to avoid unrealistic flexibility of the pyranose rings.³¹ The V_2 force constant for the internal rotation of the exocyclic CH_2OH group was taken from ref. 32. The complete parameter set of the new force field, called UNITAT, is given in Table II. "Old-fashioned" kilocalorie-per-mole (kcal/mol) energy units are used to facilitate the comparison with the original GROMOS values.

Each substance was studied only in its experimental space group. The structures were generated using the GROMOS87 force field, after which energy minimizations for a maximum of 1000 low-energy structures were performed with several force fields. In this stage, Ewald summation was used throughout, except for the C_6 terms in the force fields named HA and KOUWIJZER, in which the parameters do not follow the required combination rule and in which a cutoff of 12 Å was used.

This procedure assumes that the minima in the potential energy in the various force fields correspond to each other with only quantitative differences in energy and geometry. This is not necessarily true, and other results would have been found if the structures had been generated from scratch in each force field separately. Our computer resources were insufficient for that task.

RESULTS

Table III gives the rankings (R) and the relative energies (ΔE) of the experimental structures in the list of structures ordered to energy, with the lowest one as the reference. The average geometric deviations are also shown; they are fairly comparable between the force fields, except for UNITAT, which performs better because it was optimized for just this set of structures.

In some cases, the experimental structure was absent from the list of hypothetical structures. This

was always due to incorrect hydrogen-bonding schemes, defined as having hydroxyl torsional angles deviating more than 60° from the observed ones. For the first six structures, the same trends are observed as in our pilot study,⁹ but the details are different. Obviously, the exact results are rather sensitive to changes in the algorithm.

It was disappointing to find that GROMOS96 is not superior to GROMOS87, although the united-atom parameters have been determined carefully³³ and show a more sensible trend. On the other hand, the fairly modest changes from GROMOS87 to UNITAT produce an enormous improvement. Note that this force field was optimized for geometry only, which turns out to produce a better energy ranking as well. As already stated by Price and Wibley,⁵ if the force field does not reproduce the experimental geometry acceptably, no reliable structure prediction can be expected.

For the three all-atom force fields the one developed by Kouwijzer et al.²⁴ would be by far the best one, if it were not for the two structures BDGLOS and GLUCSE. The OPLS force field also shows two extremely large values for R and ΔE . However, the origin is different here: if only the experimental conformation is considered, all rankings are below 40, whereas the bad rankings remain unchanged for the other two force fields. This means that only the energy differences between conformations need to be improved, for instance, by adjusting some dihedral force constants. We did not attempt this for lack of data, but decided to select the OPLS force field as the most promising one of the all-atom force fields.

An essential flaw in all these force fields is the absence of explicit polarization. Because only pair potentials are present, cooperation is lacking: a favorable chain of hydrogen bonds should have a lower energy than the sum of comparable hydrogen bonds in isolation. A straightforward way to deal with this is to make the atoms polarizable, with an induced dipole that depends on the local field strength.³⁴ This formalism was included in UPACK, which made the calculations considerably slower. It is possible to include all atoms, with damping to avoid polarization catastrophes,³⁵ or to consider only intermolecular interaction. Apart from adding atomic polarizabilities, the force fields had to be reparameterized because the usual charges contain the average effect of polarization. Again, the approach was to minimize the average D , but we were unable to find an acceptable parameterization for an extension of either UNITAT or OPLS: in both cases, all rankings in the test set were

TABLE II.
Parameters for UNITAT Force Field.

Bonds	k_b	b_0		
O—CH3	600	1.424		
O—other C	600	1.413		
C—C	600	1.515		
H—O	750	1.000		
Angles	k_θ	θ_0		
C—O—C	110	110.5		
O—C—O	110	113.7		
O—C—C	120	107.5		
C—C—C	110	112.0		
H—O—C	95	107.0		
Improper Dihedrals	k_ξ	ξ_0		
Chirality at CH1	80	± 35.3		
Ring dihedral angle	10	± 57.0		
Dihedrals	V_n	n		
X—C—C—X'	1.4	3		
H—OA—C—X	0.3	3		
C—OS—C—X	0.9	3		
OA—C—C—O	0.5	2		
OA—CH2—C—OS	1.1	2		
O—C—C—C	0.1	2		
Charges: Every C—C Bond Separates Two of the Following Charge Groups:				
C—O—H	0.15, −0.50, 0.35			
CH3	0.00			
C—O—C—O—H	0.20, −0.36, 0.31, −0.50, 0.35			
C—O—C—O—CH3	0.15, −0.36, 0.40, −0.36, 0.17			
C—O—C	0.18, −0.36, 0.18			
Lennard–Jones Parameters	C_6	C_{12}		
H	0.00	0		
OA	300.0	540.6	360000	
OS	300.0	540.6	177240	
CH0	559.3		806400	
CH1	2986.6	695.9	17148000	893000
CH2	2174.4	1129.0	8444800	1700400
CH3	2121.5	1637.8	6250000	2883200

Atom types are H, OA (hydroxyl O), OS (ether O), CH0, CH1, CH2, CH3 (united C); the symbols O and C denote all subtypes. In dihedrals, X and X' denote all connected atoms, and the value of V_n is divided by the number of possibilities to define the dihedral. Second entries for Lennard–Jones parameters indicate different values for intramolecular 1···4 interactions. Mixed interaction parameters are found from the geometric mean, except OA···OS, which is the same as OA···OA and mixed interactions with OA, which are calculated as for OS. Units are kcal/mol for energy and Å for distances.

TABLE III.
Results of Force Field Comparison.

	GROMOS87		GROMOS96		UNITAT		HA		KOUWIJZER		OPLS	
	<i>R</i>	ΔE	<i>R</i>	ΔE	<i>R</i>	ΔE	<i>R</i>	ΔE	<i>R</i>	ΔE	<i>R</i>	ΔE
Substance												
ADGALA03	5	2.8	9	2.2	6	5.6	17	5.6	1	0.0	8	8.5
GLUCSA10	12	8.6	29	10.8	9	5.5	89	11.2	6	4.8	31	14.1
ADTALO10	25	9.2	6	2.7	10	2.5	42	9.7	1	0.0	57	16.5
COKBIN	3	3.9	25	9.5	23	5.2	157	16.3	1	0.0	50	10.5
BDGLOS01	77	18.2	11	10.9	51	13.0	99	14.1	197	23.7	46	14.6
GLUCSE01	192	23.8	—	—	35	10.2	—	—	—	—	5	4.9
MALTPY01	329	18.9	250	19.4	91	11.3	60	11.6	38	11.2	10	6.7
MGLUCP11	10	5.9	8	4.8	4	2.3	—	—	28	7.2	11	14.8
MEMANP11	33	11.6	44	9.4	12	4.2	2	6.6	1	0.0	1	0.0
MBDGAL02	28	10.2	9	8.3	12	7.1	—	—	14	3.6	1	0.0
IDITOL	3	5.1	4	4.3	2	0.2	296	21.8	2	0.4	91	20.0
GLUCIT01	19	8.1	14	4.2	2	1.2	192	19.9	5	1.4	191	28.0
DMANTL	4	4.7	6	5.8	1	0.0	44	11.0	13	3.2	1	0.0
DMANTL01	12	7.5	15	8.7	4	2.9	32	9.6	2	0.5	8	14.2
Average geometrical deviations												
<i>B</i>	0.024		0.027		0.017		0.020		0.024		0.020	
<i>A</i>	3.0		2.8		1.6		2.0		2.1		2.0	
<i>I</i>	2.8		3.2		1.6							
<i>T</i>	5.5		5.9		4.0		3.0		2.5		3.3	
<i>TH</i>	14		14		12		18		13		12	
<i>HB</i>	0.15		0.16		0.10		0.15		0.14		0.15	
<i>C</i>	0.40		0.34		0.29		0.49		0.41		0.45	
<i>G</i>	0.19		0.23		0.17		0.29		0.23		0.22	
<i>E</i>	4.2		4.0		3.5		6.0		4.6		3.7	
Overall <i>D</i>	4.9		5.3		3.5		5.6		4.9		4.9	

Rankings, *R*, and energy differences, ΔE (kJ/mol), refer to a search in the experimental space group for several conformations. See Table I for definitions of symbols for geometrical deviations (Units: Å, degrees).

dramatically worse than in Table III. Probably, the required changes in the force field are too drastic to be found by this simple method of parameter optimization.

Structure Generation for Large Set of Compounds

SELECTION OF STRUCTURES

To investigate the performance of UPACK in combination with the two best force fields, a larger set of structures was selected. The Cambridge Structural Database²⁸ was searched for fragments with a chain of at least four adjacent carbon atoms each carrying a C—O bond. Molecules with double bonds or containing atoms other than C, H, and O were rejected. We chose to select only

simple pyranoses and polyalcohols with, at most, eight carbon and oxygen atoms, which resulted in 83 distinct structures (of which 13 have been measured with neutron diffraction). The limitations of our program excluded 11 hydrates, 6 structures with two independent molecules in the unit cell, and 3 highly symmetric space groups (*P*32, *I*₄*a*, and *R*3). This left 39 pyranoses and 24 polyalcohols.

Only optically pure pyranoses were studied (which excluded three racemates and two substances with mirror symmetry in the free molecule). Generally three enantiomorphous space groups, viz. *P*2₁2₁2₁, *P*2₁, and *P*1, cover about 90% of such structures, and only these space groups were investigated. In the test set they occurred 27, 6, and 0 times, respectively; space group *C*2 was encountered once. All conformations that could be

constructed by considering three staggered settings about each single bond, together with the two possible chair forms of the six-membered ring, were considered. Thus, there were substances with 2, 6, 18, and 54 possible conformations; the latter type, occurring only once, was not studied. Altogether 32 pyranose structures were left.

For the polyalcohols the number of possible conformations is unmanageably large for this survey of 24 structures. Moreover, the variability in space groups is much larger than for the pyranoses. For these reasons we studied each structure only in the experimental space group and in the corresponding conformation. Thus, this exercise cannot be properly considered as an attempt at crystal structure prediction: it is mainly a test of the search procedure. The ability of the force fields to account for the intramolecular energies of various conformations is not verified in this way; nevertheless, even then, predicting the correct structure is by no means always easy due to the large number of orientations for the hydroxyl groups.

CALCULATIONS

In the first stage, the modified GROMOS87 force field just described was used without changes. The cell axes were varied from 3.5 Å to 25 Å in steps of 1.5 Å, the Euler angles were varied in steps of 20°, and the centers of gravity were varied in steps of 1 Å. A density of 1.5 g cm⁻³ was assumed throughout. For each conformation the 2000 structures with lowest energy were taken over to the second stage, except for the space group *P*2₁/*c* in which so many structures are generated that this number had to be doubled.

The investigation of possible combinations of the hydroxyl dihedral angles in the second stage

was now done with the UNITAT force field. Only a limited energy minimization was carried out to save computer time. The resulting structures within an energy window of 30 kJ/mol were used as input for the third stage. Here, the same force field was used, but this time with Ewald summation. The energy minimization was continued until the root-mean-square forces were less than 5×10^{-5} kJ/mol · Å.

After that, the procedure was continued with the OPLS force field. The atomic charges were found by the "bond increment" algorithm reported in Table IV. For the substances discussed by Damm et al.,²⁵ the correct charges are reproduced, and for other substances well-defined results are easily obtained. The other parameters from ref. 25 can be directly introduced into eq. (1), paying due attention to factors of 2 in force constants, the sign changes for V_2 values, and the factor 0.5 with which all 1 ··· 4 interactions are scaled.

Some statistics for a few space groups are given in Table V. Note how the numbers of grid points

TABLE IV.
Charges for OPLS Force Field.

Negative Side	Positive Side	Bond Increment
O (anomeric)	H	0.435
O (other)	H	0.418
O (hydroxyl)	C	0.265
O (ether)	C	0.200
C	C	0
C (anomeric)	H	0.10
C (with hydroxyl)	H	0.06
C (other)	H	0.03

Each bond carries a "bond increment," which contributes equal positive and negative charges to the two atoms involved.

TABLE V.
Some Space Group Statistics.

	<i>P</i> 1	<i>P</i> 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ / <i>c</i>
Number of molecules per cell	1	2	4	4
Dimension of the grid search	8	8	8	9
Cambridge structural database (CSD) abundance percentage ^a	5	32	54	—
CSD abundance percentage ^b	1	6	10	36
Grid points (millions)	0.1	1	2	25
CPU minutes, first stage ^c	5	15	40	600
CPU minutes, second stage ^c	10	30	60	300
CPU minutes, third stage ^c	5	10	30	200

The numbers of grid points and the computing times are averages for one conformation of various substances.

^a for optically pure substances only; ^b for all substances; ^c for a Silicon Graphics computer equipped with a 180-MHz R10000 processor.

increase with the number of molecules in the cell and with the number of parameters to be varied. The number of molecules in the cell does not influence the number of grid points in the asymmetric unit, but it does affect the total number of possibilities, because a larger cell volume allows the possibility of larger cell axes. Other symmetry aspects also play a role; for instance, the three cell axes are equivalent in $P2_12_12_1$, but not in $P2_1/c$. The computing time is dependent on the number of grid points, but large variations between substances occur due to differing numbers of accepted structures.

COMPARISON WITH EXPERIMENT

It was generally easy to determine whether or not an experimental structure was present in the list of generated structures, except for the hydroxyl torsional angles. Unless hydrogen coordinates have been determined by neutron diffraction, the possibility of errors in the published values cannot be excluded. Therefore, a disagreement between calculated and observed torsional angles may indicate either a failure of the structure prediction or an error in the experimental data. In a first analysis, a discrepancy of 60° was allowed.

For pyranoses, the rankings, R , relative energies, ΔE , and the geometrical disagreement indices, D , are given in Table VI. In comparison with Table III, the rankings have now somewhat deteriorated because two more space groups as well as inverted rings were considered. To allow for a separate discussion of intramolecular and intermolecular effects, we also report the results that would have been obtained if the experimental conformation was considered only. Allowing more conformations not only increases many rankings, but often leads to structures with lower energy; in a few cases, with a fully hydrogenated ring atom C(5), even an inverted ring conformation was predicted. This suggests that the force fields do not give an adequate description of conformational energy differences.

The results for polyalcohols are given in Table VII. As discussed previously, the rankings and energy differences must be considered as lower limits here. For iditol, mannitol, and glucitol, the data in Table III are more realistic, because a large set of conformations was studied, albeit only in one space group.

In only three cases (ZZZRCI, SUNGOX, and VOMXEA) was the experimental geometry not encountered at all, which was due to an insufficient number of

structures carried over from the first stage to the second. These failures might be remedied by changing some settings of the protocol, at the cost of a large increase in computing time.

There were four pyranoses in which the experimental geometry was only recognized if the test for the hydroxyl torsional angles was omitted. More frequently, a structure differing only in hydrogen-bonding scheme came out with a lower energy than the experimental one. However, we only started to mistrust an experimental structure if both UNITAT and OPLS force fields were in agreement about alternative hydrogen positions. Interestingly, that agreement occurred only in the four cases in which the experimental hydroxyl torsional angles could not be reproduced within 60° . For BOLNUL there were even two possibilities to choose from, which might correspond to a disordered crystal. All deviating angles are reported in Table VIII. In about half the cases, the new torsional angles were also found immediately upon energy minimization of the experimental structure, but in the other half the energy barriers were too high and a local energy minimum was retained, often with a remarkably large value of D as a further indication that something was wrong. For ALFUCO,³⁶ BOLNUL,³⁷ and CAJNAC,³⁸ the proposed hydroxyl orientations produced lower energies and D values as well as a more agreeable hydrogen-bonding scheme than the published structures. The latter showed quite unfavorable O—H \cdots O angles, and we assume that they were in error. For DACHIY,³⁹ no preference for either model could be given. Nevertheless, in all four cases, the corrected results are used in Table VI, where they are marked with a superscript "b." The only remaining instance of a large D is for BDGLOS, where the OPLS force field rotates one torsional angle about 52° , thus generating a different (and probably incorrect) hydrogen-bonding scheme.

POLYMORPHISM

In our data set there are two sets of different polymorphs of the same substance. D-Mannitol occurs in a β form (DMANTL) and in a κ form (DMANTL01), both in space group $P2_12_12_1$. Methyl β -L-arabinose is found in $P2_1$ (MBLARA01) as well as in $P2_12_12_1$ (MBLARA10, equivalent to the β -D form MBLARB and not to be confused with the α -L form MALARA10). The energy differences between the polymorphs can be found by inspection of Tables VI and VII. They are quite small except for

TABLE VI.
Crystal Structure Prediction for 32 Pyranoses.

Substance	N_{conf}	UNITAT					OPLS				
		D	$R(a)$	$R(b)$	$\Delta E(a)$	$\Delta E(b)$	D	$R(a)$	$R(b)$	$\Delta E(a)$	$\Delta E(b)$
ADGALA03	6	1.8	1	7	0.0	5.6	2.5	5	9	0.9	8.5
GLUCSA10 ^a	6	1.8	9	13	5.5	5.5	5.9	17	45	14.1	14.1
ADTALO10	6	1.7	10	15	1.7	2.5	2.1	23	68	9.7	16.5
COKBIN	6	3.2	18	28	9.2	9.2	3.9	26	65	10.8	12.7
BDGLOS01	6	3.5	39	67	13.0	13.0	12.2	43	66	14.7	14.7
GLUCSE01	6	3.5	24	48	10.2	10.2	5.2	4	5	4.9	4.9
MALTPY01 ^a	18	6.5	28	112	11.2	11.2	3.6	4	17	7.1	9.2
MGLUCP11 ^a	18	3.3	3	4	1.0	2.3	3.0	11	18	14.8	14.8
MEMANP11 ^a	18	2.7	10	17	5.3	5.3	2.7	1	1	0.0	0.0
MBDGAL02 ^a	18	2.9	11	12	7.1	7.1	2.5	1	1	0.0	0.0
ABINOS01 ^a	2	3.3	14	191	4.8	12.1	2.7	18	53	6.9	12.0
ADMHEP	18	2.7	2	18	0.2	3.7	3.0	1	8	0.0	9.9
ALFUCO ^b	2	3.7	1	1	0.0	0.0	3.4	5	6	9.6	9.6
BAXNET	18	2.9	3	11	1.0	9.5	3.7	147	529	23.2	34.1
BDGHEP	18	2.2	13	16	4.3	4.3	3.3	1	1	0.0	0.0
BECGUL	6	1.9	9	22	3.9	8.3	3.9	9	16	5.5	7.5
BOLNUL ^b	18	4.8	1	1	0.0	0.0	2.9	2	2	2.0	2.0
OR:		3.4	2	7	2.1	2.1	2.8	1	1	0.0	0.0
CAJNAC ^b	6	2.6	52	192	8.1	16.2	2.5	8	57	15.2	15.2
CELTUI	6	2.0	11	20	2.4	8.0	1.6	2	12	1.6	6.5
DACHIY ^b	6	6.2	9	9	6.9	6.9	6.2	13	32	11.4	11.4
FRUCTO02 ^a	6	2.5	1	1	0.0	0.0	3.0	8	10	10.2	10.2
LYXOSE01 ^a	2	2.6	2	2	5.1	5.1	2.9	14	17	17.3	17.3
MALARA10	6	2.1	8	16	2.8	2.8	3.3	15	37	4.2	14.0
MBLARA01	6	3.6	4	124	0.8	11.9	2.6	2	2	0.1	0.1
MBLARB	6	4.3	5	131	1.1	12.2	3.6	1	1	0.0	0.0
MDRIBP02 ^a	6	6.9	1	18	0.0	5.7	4.2	3	6	2.1	2.1
TICXUY	6	2.3	5	12	2.6	2.6	2.5	1	2	0.0	7.6
VERTER	6	2.2	1	1	0.0	0.0	3.8	3	3	0.8	0.8
XYLOBM01 ^a	6	2.2	17	17	6.5	6.5	2.8	6	11	12.4	12.4
XYLOSE03 ^a	2	3.8	44	75	9.0	12.1	2.8	3	10	5.7	7.2
ZZZIPE01	6	3.2	3	3	1.2	1.2	3.3	4	4	1.5	1.5
ZZZRCI01	6	3.3					3.0				
Average		3.2			4.1	6.5	3.6			6.6	8.9

Rankings and energy differences (kJ/mol) refer to a search in $P1$, $P2_1$, and $P2_12_1$ for: (a) the experimental conformation only; (b) N_{conf} conformations.

^a Measured by neutron diffraction.

^b Refers to adapted hydroxyl dihedral angles, see Table VIII. For one dihedral in BOLNUL two values are acceptable.

the OPLS value for mannitol (14.2 kJ/mol), which is nevertheless within the accepted range of less than 20 kJ/mol.³

For three optically pure substances the crystal structures of their racemates are also known. Upon energy minimization it was found that the relative energies of the racemic crystals varied between -8 kJ/mol and $+12$ kJ/mol. In two of the three cases positive values occurred, which should not take place if energies were the only determining factor for crystallization: from a racemic solution

both racemic crystals and conglomerates can be formed.⁴⁰

Within a certain energy window every generated structure corresponds in principle to a potential polymorph. In reality, the number of observed polymorphs is comparatively low. One explanation can be that many of the hypothetical structures differ only with respect to the hydrogen-bond network. Assuming that, in a real crystal, there is enough opportunity for internal rotation of hydroxyl groups around the C—O bonds, these

TABLE VII.
Crystal Structure Generation for 24 Polyalcohols.

Substance	Space Group	UNITAT			OPLS		
		<i>D</i>	<i>R</i>	ΔE	<i>D</i>	<i>R</i>	ΔE
IDITOL	$P2_1$	5.2	1	0.0	5.3	17	8.0
GLUCIT01 ^a	$P2_12_12_1$	2.8	1	0.0	2.9	27	13.0
DMANTL	$P2_12_12_1$	3.1	1	0.0	3.5	1	0.0
DMANTL01	$P2_12_12_1$	3.1	3	2.9	4.5	7	14.2
ALITOL01	$P2_1/c^b$	7.0	12	7.2	4.2	1	0.0
ARABOL	$Pna2_1$	2.4	1	0.0	3.6	1	0.0
DLMANT	$Pna2_1$	1.9	1	0.0	6.3	5	3.2
EPINOS	$P2_1/c$	1.9	59	10.9	2.6	1	0.0
FAJLOR	$P2_12_12$	3.7	5	6.5	4.8	1	0.0
FOPKOK	$P2_1$	2.6	5	4.8	3.1	8	11.8
GALACT	$P2_1/c$	3.4	4	1.9	3.0	3	1.4
JINDUF	$P2_1$	1.9	6	5.1	3.2	1	0.0
JOJZOX	$P2_1$	4.5	7	5.7	4.9	6	2.6
LEDNEN	$P2_12_12$	2.7	1	0.0	2.9	1	0.0
LEDNIR	$P2_12_12_1$	2.4	4	1.2	3.3	4	3.1
MUINOS	$P2_1/c$	1.2	3	1.7	2.2	1	0.0
RIBTOL	$P2_1/c$	5.8	1	0.0	5.8	7	8.2
SUNHAK	$P\bar{1}$	4.6	2	0.8	4.2	5	3.3
SUNGOX	$Pbca$	2.9			4.6		
TEBCUY	Cc	2.3	1	0.0	3.7	1	0.0
VOMXEA	$P2_1/c$	3.3			4.2		
VOXXOV	$P2_12_12_1$	2.6	1	0.0	4.1	1	0.0
XYLTOL	$P2_12_12_1$	2.9	1	0.0	3.0	3	2.1
YEPNOW	$P\bar{1}^c$	1.4	4	7.6	2.3	1	0.0
Average		3.2		2.6	3.8		3.2

Rankings and energy differences (kJ/mol) refer to the experimental conformation in the experimental space group only.

^a Measured by neutron diffraction.

^b $P2_1/c$ with molecule on inversion center, treated as $P2_1$.

^c $P\bar{1}$ with molecule on inversion center, treated as $P1$.

TABLE VIII.
Proposed Alternatives for Published Hydroxyl Torsional Angles.

Substance	Dihedral	Value (°)		O···O (Å)		O—H···O (°)	
		X	P	X	P	X	P
ALFUCO	H2—O2—C2—C3	122	62	2.74	2.68	96	160
BOLNUL	H4—O2—C2—C3	−142	47	2.76	2.68	115	165
	H10—O4—C4—C5	151	−143	2.93	2.79	98	167
	or:		65		2.77		174
CAJNAC	H20—O2—C2—C3	−132	51	2.78	2.64	110	165
DACHYI	H1—O1—C1—C2	170	−38	2.67	2.70	169	158
	H6—O6—C6—C5	99	−73	2.71	2.67	171	170

The values calculated from X-ray measurements in the literature (X) are compared with the values proposed (P) by calculations with the OPLS force field. The differences in other dihedrals are less than 40°.

structures are in fact identical. The numbers of remaining structures vary enormously between the different substances. The averages for the entries in Table VI are shown in Figure 1 as a function of their relative energy. The difference between the three space groups follows the number of grid points in Table V; $P2_1/c$ would be even more prolific. Less obvious is the appreciable difference between the two force fields studied, with OPLS giving larger energy differences than UNITAT.

It is quite possible that several of these structures can still interconvert if enough thermal energy is available. This subject is outside the scope of the present study, but it is unlikely that the numbers of distinct hypothetical structures will ever become very low. This suggests an enormous potential for polymorphism, which, theoretically, should be the rule rather than the exception. Indeed, according to McCrone,⁴¹ the number of polymorphs of a given substance is proportional to the effort that one is prepared to put into crystallization experiments. Considering that Sarma and Desiraju⁴² have identified only 3.5% of the entries in the CSD as polymorphic, a fertile field of research should still be open here.

Discussion

Of course, in a real structure prediction no experimental crystallographic data should be used. Fortunately, one can always exclude nonenantiomorphous space groups for optically pure substances. For this study, we did not consider the possibility that a hydrate may crystallize from aqueous solution. As discussed earlier, for polyalcohols the problem is still too complex for a complete analysis on a routine basis. Without crystallographic knowledge our test set of optically pure

anhydrous pyranoses would have contained 36 substances, for which UPACK structure prediction would fail in five cases: two because there are two molecules in the asymmetric unit; two because of exceptional space groups ($P32$ and $C2$); and one because the correct structure was wrongly discarded after the first stage. In the other 31 cases, the remaining problem is to select the observable polymorph(s) out of the list of hypothetical structures.

From Tables III and VI, there is little guidance to choose between the two force fields studied. It is of interest to note that averaging the two sets of energy values leads to much better rankings than either set alone, but there seems to be little future in such an opportunistic mixing of results from two essentially different force fields. The OPLS force field suffers from one spectacular failure (BAXNET); unlike for IDITOL and GLUCIT (Tables III and VII), there is hardly any improvement if only the experimental conformation is considered. Such a result is unacceptable, even for a force field of the simple form used here, and we are planning work on further improvement.

The UNITAT force field performed remarkably well, taking into account that it uses the primitive concept of united atoms. It may be argued that this conclusion is not fair because the force field was optimized on the same structures that are being predicted. However, the optimization used only experimental geometric data and no energetic information, whereas the force field was tested for its ability to produce acceptable relative energies. A minor consequence of the optimization procedure is that exactly the same geometries would have been produced if all energies would have been multiplied with an arbitrary factor. If we wish to bring the average lattice energies of our molecules on the same scale as the OPLS force field

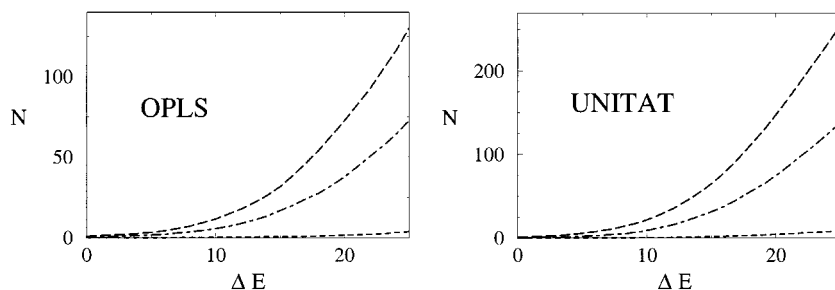


FIGURE 1. Average numbers (N) of hypothetical structures, as functions of their relative energies, ΔE (kJ/mol), for the pyranoses reported in Table VI. Short-dashed line: $P1$; dot-dashed line: $P2_1$; long-dashed line: $P2_12_1$. The UNITAT energies have been scaled with a factor 1.2 to make them roughly comparable to the OPLS values.

(which does contain experimental information on energetics), that factor should be about 1.2. This should be kept in mind when comparing the force fields; in fact, this factor is included in Figure 1 but not in the tables.

Regardless, the stage was reached where the average relative energies of the experimental structures (ΔE in Table VI) were well below 10 kJ/mol in a large set of test molecules. This suggests that typical energy differences between polymorphs cannot be much larger than that, at least for comparatively small molecules like the ones studied here. However, in a recent study, an extremely large energy difference (20 to 25 kJ/mol) was calculated between two polymorphs of *o*-acetamidobenzamide.⁴³ Such a value must be considered exceptional in view of the results obtained here and in other crystal structure generations.

The rankings and energy differences for the six hexapyranoses considerably improved with respect to the previous investigation (the first six structures of Tables III and VI correspond in that order to Table 3 of ref. 9, where rankings of a few hundred were found). On the whole, only three experimental structures were missed and few rankings larger than 100 occurred. This means that the number of candidates for observable structures is becoming so low that each can be subjected to much more sophisticated energy calculations. We are presently working on a force field that combines atomic multipole moments from quantum-chemical *ab initio* calculations with intermolecular polarization and a reliable intramolecular force field. This should limit the number of possible polymorphs even further. In that case one could also perform molecular dynamics calculations on all predicted structures. On one hand, some structures might interconvert, and thus be shown to be identical⁷ and, on the other hand, this would give a rough indication of the temperature effects on the energy. Entropy differences could be estimated from lattice vibrations, and that would open the road to a real *ab initio* calculation of polymorphic transitions.

If not only the energies, but also the geometries of the hypothetical structures could be further improved, more reliable powder diffraction diagrams could be calculated for comparison with experimental data in cases in which no single crystal data can be obtained.

As noted earlier,⁴⁴ even without knowing the correct structure it is possible to estimate the density and lattice energy of any (not too complex) substance according to a given force field. The

reason is that, for these properties, the polymorph with lowest energy will be comparable to the real structure(s) within the accuracy of the force field. Thus, no experimental information is needed for such a calculation.

The UPACK program package is under constant development. The programs were developed with routines from the GROMOS molecular dynamics simulation program package¹⁰ as basis, and we are grateful to Professor W. F. van Gunsteren for his permission to do so. They are written in standard FORTRAN77 without any graphical interface, so they should be easy to install, but their use requires some familiarity with file manipulation. The current source codes and program manual can be obtained upon application by e-mail to vaneyck@chem.uu.nl. The coordinates of specific hypothetical low-energy structures obtained in the present work are also available on request.

References

1. Gdanitz, R. J. In: Gavezzotti, A., ed. *Theoretical Aspects and Computer Modeling of the Molecular Solid State*; Wiley: Chichester, UK, 1997; pp. 185–201.
2. Verwer, P.; Leusen, F. J. J. In: Lipkowitz, K. B.; Boyd, D. B., eds. *Reviews in Computational Chemistry*; Wiley-VCH: New York, 1998; pp. 327–365.
3. Gavezzotti, A. *Acc Chem Res* 1994, 27, 309.
4. van Eijck, B. P.; Spek, A. L.; Mooij, W. T. M.; Kroon, J. *Acta Crystallogr* 1998, B54, 291.
5. Price, S. L.; Wibley, K. S. *J Phys Chem* 1997, A101, 2198.
6. Coombes, D. S.; Nagi, G. K.; Price, S. L. *Chem Phys Lett* 1997, 265, 532.
7. Mooij, W. T. M.; van Eijck, B. P.; Price, S. L.; Verwer, P.; Kroon, J. *J Comput Chem* 1998, 19, 459.
8. Karfunkel, H. R.; Gdanitz, R. J. *J Comput Chem* 1992, 13, 1171.
9. van Eijck, B. P.; Mooij, W. T. M.; Kroon, J. *Acta Crystallogr* 1995, B51, 99.
10. van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation Package (GROMOS)*, University of Groningen, 1987.
11. Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes, the Art of Scientific Computing*; Cambridge University Press: Cambridge, UK, 1989.
12. van Eijck, B. P.; Kroon, J. *J Comput Chem* 1997, 18, 1036.
13. Pérez, S.; Imberty, A.; Engelsens, S. B. *Carbohydr Res.* In press.
14. Glennon, T. M.; Zheng, Y.-J.; Legrand, S. M.; Shutzberg, B. A.; Merz, K. M. *J Comput Chem* 1994, 15, 1019.
15. Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J Phys Chem* 1995, 99, 3832.
16. Senderowitz, H.; Still, W. C. *J Org Chem* 1997, 62, 1427.

17. Glennon, T. M.; Merz, K. M. *J Mol Struct (Theochem)* 1997, 395–396, 157.
18. Coombes, D. S.; Price, S. L.; Willock, D. J.; Leslie, M. *J Phys Chem* 1996, 100, 7352.
19. Ma, B.; Schaefer, H. F.; Allinger, N. L. *J Am Chem Soc* 1998, 120, 3411.
20. Hwang, M.-J.; Ni, X.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *Biopolymers* 1998, 45, 435.
21. van Gunsteren, W. F., et al. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Hochschulverlag: Zürich, 1996.
22. Ott, K.-H.; Meyer, B. *J Comput Chem* 1996, 17, 1068.
23. Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr Res* 1988, 180, 207.
24. Kouwijzer, M. L. C. E.; van Eijck, B. P.; Kooijman, H.; Kroon, J. *Acta Crystallogr* 1995, B51, 209.
25. Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. *J Comput Chem* 1997, 18, 1955.
26. Rasmussen, K. *J Mol Struct (Theochem)* 1997, 395–396, 91.
27. Fabricius, J.; Engelsen, S. B.; Rasmussen, K. *J Carbohydr Chem* 1997, 16, 751.
28. Allen, F. H.; Kennard, O. *Chem Design Automat News* 1993, 8, 31.
29. Jeffrey, G. A.; Kim, H. S. *Carbohydr Res* 1970, 14, 207.
30. Schouten, A.; Kanters, J. A.; Kroon, J.; Comini, S.; Looten, P.; Mathlouthi, M. *Carbohydr Res* 1988, 312, 131.
31. Spieser, S.; Kroon-Batenburg, L. M. J. Private communication.
32. Klewinghaus, P.; van Eijck, B. P.; Kouwijzer, M. L. C. E.; Kroon, J. *J Molec Struct (Theochem)* 1997, 395–396, 289.
33. Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J Comput Chem* 1998, 19, 535.
34. van Eijck, B. P.; Kroon-Batenburg, L. M. J.; Kroon, J. In: Gavezzotti, A., ed. *Theoretical Aspects and Computer Modeling of the Molecular Solid State*; Wiley: Chichester, UK, 1997; pp. 99–146.
35. Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. *J Phys Chem* 1994, 98, 4180.
36. Longchambon, F.; Ohannessian, J.; Avenel, D.; Neuman, A. *Acta Crystallogr* 1975, B31, 2623.
37. Sheldrick, B.; Akrigg, D. *Acta Crystallogr* 1983, C39, 315.
38. Neuman, A.; Gillier, H. *Acta Crystallogr* 1983, C39, 1382.
39. Puliti, R.; Mattia, C. A.; Barone, G. *Carbohydr Res* 1984, 135, 47.
40. Brock, C. P.; Schweizer, W. B.; Dunitz, J. D. *J Am Chem Soc* 1991, 113, 9811.
41. McCrone, W. C. In: Fox, D.; Labes, M. M.; Weissberger, A., eds. *Physics and Chemistry of the Organic Solid State*, Vol. II; Interscience: New York, 1965, pp. 725–767.
42. Sarma, J. A. R. P.; Desiraju, G. R. In: Seddon, K. R.; Zaworotko, M. J., eds. *Crystal Engineering. The Design and Application of Functional Solids*; Kluwer: Dordrecht, 1998.
43. Buttar, D.; Charlton, M. H.; Docherty, R.; Starbuck, J. *J Chem Soc Perkin Trans* 1998, 763.
44. Gavezzotti, A. *Faraday Disc* 1997, 106, 63.